

**REMARKS**

Reconsideration and allowance are requested. New claims 13-36 were previously presented in this application. Claim 23 is amended in this response.

**Rejection of Claim 23 Under Section 112**

The Examiner rejects claim 23 under Section 112. Applicants have amended claim 23 to recite a "numeric understanding processor". This amendment does not narrow the scope of the claim but merely makes a minor change in the claim language to conform the language to the specification. Therefore, Applicants respectfully request that the Examiner withdraw this rejection of claim 23.

**Rejection of Claim 28 Under 35 U.S.C. 102**

The Examiner rejects claim 28 under Section 102 as being unpatentable over U.S. Patent No. 5,970,449 to Alleva et al. ("Alleva et al."). Applicants respectfully traverse this rejection and submit that claim 28 is patentable over Alleva et al.

Claim 28 recites a system comprising a speech recognition processor that receives unconstrained input speech and outputs a string of words. The speech recognition processor is based on a numeric language that represents a subset of a vocabulary, the subset including a set of words identified as being relevant for interpreting and understanding number strings. The Examiner asserts that Alleva et al. disclose this limitation at various places. (Col. 3, line 18 to col. 4, line 6; Abstract; Figure 1, element 32; Figure 9, element 132; col. 1, lines 56 - 62; col. 6, lines 14 - 17 and 40 - 42; col. 5, lines 62 - 65 and col. 6, lines 32 - 64). Applicants submit that as explained below, the text normalization that transforms text into digits disclosed in Alleva et al. does not correspond or disclose any limitation in claim 28.

Alleva et al. teach a text normalizer that takes the output from the speech recognizer and produces text that is more familiar to people who will read the text. Applicants will address each portion of Alleva et al. referenced by the Examiner to explain that the references does not disclose a speech recognition processor that is based on a numeric language that represents a subset of a vocabulary.

**Col. 3, line 18 - col. 4, line 6 and Figure 1, element 32**

This portion of Alleva et al. discusses FIG. 1 which discloses a system embodiment of their invention. In that figure, the Examiner cites element 32 as subject matter that corresponds to claim 28. However, feature element 32 is the speech recognizer and not the

text normalizer, which is feature 38. As is clear from FIG. 1, the text normalizer 38 is a post-processing component to the speech recognizer 32. The text normalizer 38 processes the text produced by the speech recognizer 32. The Examiner blends the functions of the text normalizer (i.e., converting words like "twenty" into numbers like "20" using a number rule) with the function of the speech recognizer to reject claim 28. However, this blending of functions is never taught or suggested by Alleva et al. The text normalizer is always a separate component to the speech recognizer and always simple operates on received text. By combining the text normalizer functions with the speech recognition functions, the Applicants respectfully submit that the Examiner is inappropriately altering the teachings of Alleva et al. to attempt to match the reference to the claim limitations.

Furthermore, the text normalizer has nothing to do with the speech recognition process. The speech recognition processor is based on a numeric language of claim 28 which includes a set of words that are identified as being relevant for interpreting and understanding number strings. The text normalizer of Alleva et al. does not perform speech recognition and only converts received text into numbers of symbols. Therefore, at its core, the text normalizer of Alleva et al. fails to perform any speech recognition function and cannot correspond to the speech recognition processor of claim 28.

#### Abstract

The Abstract introduces the text normalizer but, as discussed above, makes clear that the text normalizer is not a speech recognizer because the text normalizer "normalizes text that is output from a speech recognizer." The context free-grammar discussed in the Abstract allows the text normalizer to normalize the received text according to whatever rule is being applied. However, there is nothing in the Abstract which teaches that the text normalizer performs any speech recognition on spoken number strings as is required in claim 28.

#### Figure 9, element 132 [sic]

The Examiner cites element 132 of figure 9 but there is no element 132. Applicants assume the Examiner means the number rule 132 or the currency rule 130. As discussed above, however, these rules are applied by the text normalizer and not a speech recognizer. Therefore, they are not disclosed as having any relevance to speech recognition and do not therefore correspond to the limitations of claim 28.

Col. 1, lines 56 - 62

This portion of Alleva et al. merely discusses the text normalizer function. The substance of which is discussed above in that it does not relate to speech recognition.

Col. 6, lines 14 - 17 and 40 - 42

This portion of Alleva et al. discloses the rules for the text normalizer to substitute digits for text received from the speech recognizer. For example, the number "1" is substituted for the text "one." This portion is similar to FIG. 9 that shows how the number, currency and digit rule area applied. However, rather than teaching speech recognition techniques, there are post-speech recognition processing rules applied by the text normalizer. They are designed to operate on received text and have nothing to do with speech recognition. The entire purpose of these rules is to substitute more understandable symbols or numbers in the place of received text.

Col. 5, lines 62 - 65

This portion of Alleva et al. merely describes the capitalization section 66 shown in FIG. 4 which includes a rule for the received text to capitalize certain letters to make it more readable for a human. This portion fails to disclose features of claim 28 for the same reason as the other portions cited by the Examiner, namely that the functions of the text normalizer have nothing to do with speech recognition as is recited in claim 28.

Col. 6, lines 32 - 64

This portion of Alleva et al. further discusses other rules applied by the text normalizer. These rules include such things as how to deal with fractions, acronyms, names of cities and states etc. In each case, however, the particular rule is applied in the context of the text normalization process. In this regard, each rule is applied to received *text* from the speech recognizer. There is no instance where the rules are applied as part of a speech recognizer. Therefore, this portion of Alleva et al. fails to disclose the speech recognizer of claim 28.

Therefore, Applicants have explained that each portion of Alleva et al. cited by the Examiner teaches a process that occurs *after* speech recognition and has nothing to do with the technical process of speech recognition. The speech recognition processor of claim 28 is based on a numeric language that represents a subset of a vocabulary, the subset including a set of words identified as being relevant for interpreting and understanding number strings.

While the text normalizer of Alleva et al. relates to numbers and text strings, it is consistently taught as a process that occurs after speech recognition and only acts on text. Since the text normalizer of Alleva et al. does not match or correspond to any speech recognition process, Applicants submit that claim 28 is not anticipated by Alleva et al. and is therefore patentable and in condition for allowance.

**Rejection of Claims 17 - 19, 21 - 27, 29 - 34 and 36 Under Section 103**

The Examiner rejects claims 17 - 19, 21 - 27, 29 - 34 and 36 under Section 103 as being unpatentable over Alleva et al. and U.S. Patent No. 5,613,037 to Sukkar ("Sukkar"). Applicants respectfully traverse this rejection and submit that these claims are patentable.

Claim 17 is discussed first. This claim recites a speech recognition method that comprises performing a speech recognition process on the received speech signal to produce speech recognition results, the speech recognition process being based on a set of acoustical models that has been defined for a numeric language, wherein the numeric language includes a subset of a vocabulary. The subset of the vocabulary includes words that identify digits in number strings and words that enable the interpretation and understanding of number strings. The Examiner asserts that Alleva et al. teach a speech recognition processor that produces textual output such as "ten cents" and "four o'clock in the afternoon". While the speech processor of Alleva et al. is shown as recognizing words such as "ten" and "four" from input speech, Applicants submit that the speech processor of Alleva et al. does not match the method of claim 17. Claim 17 requires the speech recognition process to be "based on a set of acoustical models that has been defined for a numeric language." It is clear that the speech recognizer in Alleva et al. may interpret some numbers but they also interpret other words as well. There is no suggestion or teaching in Alleva et al. that their speech recognition process is based on acoustic models as recited in claim 17.

In this regard, Alleva et al. teach away from the invention recited in claim 17 in that their recognizer is more of a general speech recognizer and not based on acoustic models that have been defined for a numeric language. Therefore, Applicants submit that claim 17 is patentable.

Further, the Examiner notes that Alleva et al. fail to disclose a system comprising acoustic models utilized by the speech recognition processor. The Examiner combines Alleva et al. with Sukkar to fill in the missing disclosure of Alleva et al. and reject claim 17. Applicants traverse this combination of references to reject this claim.

To establish a *prima facie* case of obviousness, the Examiner must meet three criteria. First, there must be some motivation or suggestion, either in the references themselves, or in the knowledge generally available to one of ordinary skill in the art, to combine the references. Second, there must be a reasonable expectation of success, and finally, the prior art references must teach or suggest all the claim limitations. The Examiner bears the initial burden of providing some suggestion of the desirability of doing what the inventor has done. "To support the conclusion that the claimed invention is directed to obvious subject matter, either the references must expressly or impliedly suggest the claimed invention or the examiner must present a convincing line of reasoning as to why the artisan would have found the claimed invention to have been obvious in light of the teachings of the references." MPEP 2142.

If a proposed modification would render the prior art invention being modified unsatisfactory for its intended purposes, then there is no suggestion or motivation to make the proposed modification. *In re Gordon*, 733 F.2d 900, 221 USPQ 1125 (Fed. Cir. 1984). Further, if the proposed modification of the prior art would change the principle operation of the prior art invention being modified, then the teaching of the reference is not sufficient to render the claims *prima facie* obvious. *In re Ratti*, 270 F.2d 810, 123 USPQ 349 (CCPA 1959). The principles outlined in both these cases are applicable here.

There is no motivation or suggestion to combine Alleva et al. with Sukkar. Alleva et al. teach a text normalization process for the purpose of producing textual output that is not awkward but is familiar to human recipients. Col. 1, lines 19 - 32. The text normalizer operates on the text obtained from the speech recognizer that includes sentences spoken by the user as shown in FIG. 9. In contrast, the purpose of Sukkar is to provide a system that receives speech and removes portions of the received speech that are not digits for the purpose of systems such as telephone systems and credit card applications. Phrases that include non-digit language are also rejected. Col. 2, lines 38 - 42. In the summary of the invention portion, Sukkar explains that received speech that includes digits and other non-digit speech is rejected by the recognizer. The entire phrase may be rejected. Sukkar explains that a recognizer and a filler model recognizer are used to analyze the speech and only pass speech containing digits to the network for processing. The purpose of Sukkar's invention is to enable an automated system to process numbers from credit cards or phone numbers, etc. by receiving a clean set of digits from the speech recognizer. Col. 3, line 19 - col. 4, line 22. Cleaning up the speech to only include digits of a high degree of reliability removes any requirement for a live person to be on a call to interpret numbers from a phone

number or a credit card. FIG. 1 illustrates the speech recognizer 122 and processor 118 connected to a toll switch for processing the numbers. There is no mention or suggestion in Sukkar that a human would be the recipient of the speech converted to text.

With this background, Applicants submit that there is no suggestion to combine Alleva et al. with Sukkar because Alleva et al.'s focus and objective is to make speech converted to text readable for humans. In contrast, Sukkar strips away non-digit text from speech and produces only digits for use by a computer network for processing. Alleva et al.'s invention would be useless if blended with Sukkar because no human reads the text produced by speech recognition in Sukkar. Further, Sukkar's invention would be useless if blended into Alleva et al.'s invention because that would remove all phrases with non-digit words contained therein from the speech recognizer. Such output would not be understandable or useable by a human reader. For example, there would be no sentences such as the one shown in FIG. 9 of Alleva et al.: "five chickens at twenty cents each." The output of that sentence would be rejected if Sukkar's speech recognizer were used because there are non-digit words in the sentence that would cause the confidence score would be less than the digit threshold and the entire string would be rejected. See Abstract of Sukkar.

As can be seen from the above explanation, for the blending of Sukkar with Alleva et al. to work, there would have to be modifications and changes in the fundamental principles of operation of one or both of the references to such an extent that it is impossible to combine them. Such dramatic modifications to these basic principles of operation only lead to the conclusion that they cannot be combined to reject claim 17. Indeed, since Sukkar teaches rejecting such phrases, Sukkar actually teaches away from any such combination with Alleva et al. which seeks to provide all the spoken phrases to a human for reading.

Therefore, there cannot be any motivation or suggestion to combine Alleva et al. with Sukkar to reject claim 17. Applicants respectfully request withdrawal of the rejection of claim 17 and submit that this claim is in condition for allowance.

Claims 18 - 19 and 21 - 27 each depend from claim 17 and recite further limitations therefrom. Accordingly, Applicants submit that these claims are patentable as well inasmuch as the references should not be combined to reject the claims and event if combined, the combination fails to teach each claim limitation.

Similarly, since these two references cannot be combined, Applicants submit that claims 29 - 34 and 36 are patentable as well and are in condition for allowance.

**Rejection of Claims 13, 16, 20 and 35 Under Section 103**

The Examiner rejects claims 13, 16, 20 and 35 under section 103 as being unpatentable over Alleva et al., Sukkar and further in view of U.S. Patent No. 5,937,384 to Huang et al. ("Huang et al."). Applicants traverse these rejections and submit that these references should not be combined to reject the claims.


As set forth above, Sukkar and Alleva et al. teach away from any combination. It is well established that there is no suggestion or motivation to combine Alleva et al. with Sukkar. Therefore, claims 13, 16, 20 and 35 are patentable and in condition for allowance since the two primary references cannot be legally combined to arrive at the inventions recited in these claims.

**CONCLUSION**

Having addressed the rejection of claims, applicant respectfully submits that the subject application is in condition for allowance and a Notice to that effect is earnestly solicited.

Respectfully submitted,

Date: May 12, 2004  
Correspondence Address:  
Samuel H. Dworetsky  
AT&T Corp.  
Room 2A-207  
One AT&T Way  
Bedminster, NJ 07921

  
Thomas M. Isaacson  
Attorney for Applicants  
Reg. No. 44,166  
Phone: 410-414-3056  
Fax No.: 410-510-1433